MI と DX
マテリアル研究 DX プラットフォーム WAVEBASE

（トヨタ自動車株式会社）庄司　哲也

Material Informatics and Digital Transformation, WAVEBASE: Material R&D DX Platform

(Toyota Motor Corporation) Tetsuya Shoji

Recent rapid evolution of computational power and generalization of AI technology enables information extraction from variety types of data. Commonly, material researchers and engineers tend to recognize AI technology as tools for machine learning in their material R&D scene. But most important point of material informatics is how we become able to access meaningful descriptor buried in high dimensional data which is difficult to quantify. Moreover, material data have multi-scale and multi-modal problem. This fact makes data utilization more difficult. In order to get over this situation, we focused on wave number "q-space" expression for handling multi scale structural data, as shown in **Figure 1**. Using q-space expression, we can connect from atomistic diffraction information to structural scattering information based on principle of diffraction and scattering. Moreover, small angle scattering information is almost equivalent to power spectrum of microscope image. Based on this fact, we can recognize that we can express multi-scale structural information in one curved line which is consist of scattering, or power spectrum, and diffraction[1]. Using superposition principle of wave, we can aware that all structural information is overlapped in this curved line.

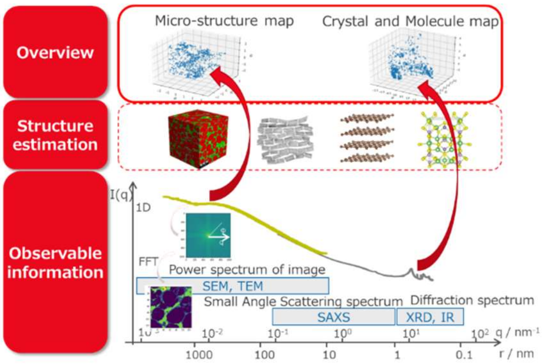

Figure 1　q-space expression of multiscale expression and corresponding measurement.

To extract information from multiscale dataset, we utilize dimensionality reduction technique for high dimensional multiscale data. For example, we start with 35 different composition and nanostructure Light Rare Earth alloyed nano structured $Nd_2Fe_{14}B$ rare-earth magnet samples and measure X-ray diffraction (XRD)[2]. One XRD data consists of 3500 dimensions. Commonly known that XRD spectrum consist of diffraction information from atomistic configuration and small angle scattering information from nano

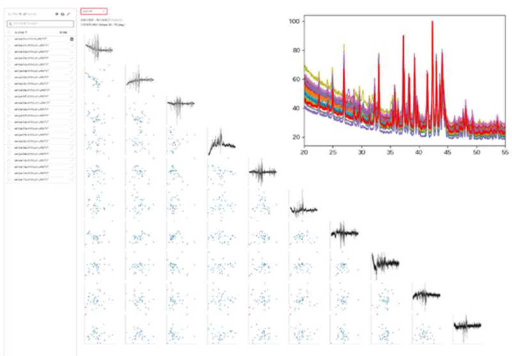

Figure 2　35 XRD data of Nd reduced rare earth magnet and 10 decomposed vectors by PCA analysis

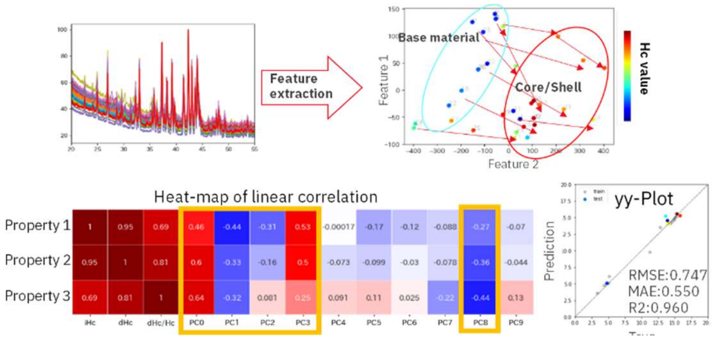

Figure 3 Correlation of properties and PC coefficient.

and microstructure. To extract feature from these 35 XRD data, we use PCA, principal component analysis, to decompose 10 independent vectors as shown in **Figure 2**. In this case cumulative contribution ratio is more than 0.95 with these 10 vectors. Looking at decomposed vector with knowledge of diffraction and small angle scattering, one can understand meaning of each vector. In this case, 1^{st} principal component (PC) corresponds to difference of scattering by nanostructure and part of peak shape change and 2^{nd} PC corresponds to peak shift of $RE_2Fe_{14}B$ phase, this means that change of lattice parameter by alloying Light Rare Earth element in $Nd_2Fe_{14}B$ phase. 3^{rd} PC represent change of peak width. Then we use coefficient of each PC as descriptor for magnetic performance machine learning model. Coefficient of PC seems to work well as good descriptor as shown in **Figure 3**. This is only one example, but we already confirmed this like analysis may hold in other materials, such as catalyst, battery materials, lubricant oil etc., and other spectrum data, such as FT-IR or GC-MS etc.

Based on our achievement of material informatics research include feature extraction by dimensionality reduction, we deployed Data storage and Data analysis platform “WAVEBASE” on AWS and started Material R&D DX support service business [3]. Using this system, users can store their data on same place in their limited community, or company and can extract feature information from proper dataset and analyze correlation against objective parameters, or target properties, by machine learning algorithms without coding. Furthermore, SINET6, high speed data transfer network, will become in service, massive data from quantum beam facilities, such as SPring-8, become able to transfer to AWS immediately for analyze measured data on site. IT environment evolution may lead transformation of analytics of massive data and efficient usage of beam time.

Our attempt is just one of option for users. This like SaaS business will become mainstream for data analytics in material R&D, I think. Most important thing for engineers and researchers is understand underlying theory, not only material itself but measurement principles and data analytics.

Keywords : Material informatics; Digital transformation; feature extraction; Dimensionality reduction;

1) S. Koizumi et al; J. Appl. Cryst. (2019). 52, 783-790
2) A. Kato et al; Matere. (2021) 60, 1, 57-59
3) URL: https://www.toyota.co.jp/wavebase/